A

PATENT
W&C REF.: 1103326-0206

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

"Express Mail" Label No EJ064077753US.

Date of Deposit October 13, 1999 I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the. Assistant Commissioner for Patents Washington, D.C 20231

Jermaine Peebles
(Type or print name of person mailing paper or fee)

Jermaine Peebles
(Signature of person mailing paper or fee)

Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

Sir:

This is a request for the filing of a continuation application under 37 CFR 1.53 (b) of pending prior application Serial No. **08/624,398,** filed on **April 4, 1996,** entitled:

DNA MOLECULES FOR EXPRESSION OF POLYPEPTIDES

for: (inventor) Goutam Das

1.(**X**) Enclosed is a copy of the prior application as originally filed.

2.( ) Small entity status of this application under 37 CFR 1.9 and 1.27 has been established by a verified statement previously submitted in USSN .

3.( ) A verified statement to establish small entity status under 37 CFR 1.9 and 1.27 is enclosed.

4.(X) The filing fee is calculated below:

FOR	(Col. 1) NO. FILED	(Col. 2) NO. EXTRA	SMALL ENTITY RATE	SMALL ENTITY FEE	or	OTHER THAN A SMALL ENTITY RATE	OTHER THAN A SMALL ENTITY FEE
Basic Fee	/////////	/////////	////	$380	or	////	$760
Tot. Claims	-20 = *		x 9 =		or	x18 =	
Ind. Claims	1 -3 = *		x39 =		or	x78 =	
(X) Multiple Dependent Claim Presented			+130 =		or	+260 =	260
		TOTAL=			or		$1020

* If the difference in Col. 1 is less than zero, enter "0" in Col 2.

5a(X) A petition for extension of time for three (3) months to respond to the office action in parent application USSN 08/624,398 has been filed. A copy is enclosed.

Or

5b.( ) In the event that an extension of time is required, this conditional petition is being made to provide for the possibility that applicant has inadvertently overlooked the need for a petition and fees for extension of time.

5c(X) TOTAL FEE DUE:

Filing fees $1020
Extension fee (if any) $________

TOTAL FEE DUE $1020

( ) A check in the amount of $1020 is enclosed.

5d (X) **THE FILING FEE IS NOT ENCLOSED**.

( ) The Commissioner is hereby authorized to charge the filing fee, excess claims fee (if applicable), excess independent claims fee (if applicable), and multiple dependent claims fee (if applicable) to Deposit Account No. 23-1703.

6. ( ) The Commissioner is hereby authorized to charge any additional filing fees required under 37 CFR 1.16 and 1.17 associated with this communication or credit any overpayment to Deposit Account No. 23-1703. Two copies of this sheet are enclosed.

7.(**X**) A Preliminary Amendment is enclosed (regarding the sequence listing). Also enclosed is the computer readable copy of the sequence listing (diskette) and a hard copy of same.

7a.( ) Please cancel claims .

8.(**X**) Amend the specification by inserting before the first line the sentence:

This application is a continuation of application Serial No. **08/624,398,** filed on April 4, 1996, which is a 371 of PCT/SE96/00318, filed March 12, 1996.

9a. ( ) Transfer the drawings from the prior application to this application and abandon said prior application as of the filing date accorded this application.

9b.( ) Two sheets of drawings are enclosed.

10.(**X**) The prior application is assigned to **Astra Aktiebolag.**

11.(**X**) a. (**X**) The Declaration and Power of Attorney appears in the original papers of parent application Ser. No.08/624,398, filed April 4, 1996. A copy of that Declaration/Power of Attorney is enclosed.

b. ( ) A copy of the Revocation and New Power of Attorney in the prior application is enclosed.

c. ( ) Since the Power does not appear in the original papers, a copy of the Power in the prior application is enclosed.

d. ( ) Recognize as associate attorneys:

____

(Name, Reg. No. and Address)

______________________________.

12.(**X**) Applicant claims priority in this application under 35 USC 119 of Indian Appln. No. 351/MAS/95, filed March 23, 1995, and Swedish Application No. 9501939-4, filed May 24, 1995. The certified copies were filed in International Application PCT/SE96/00318.

13(**X**) A second duplicate copy of this letter is enclosed for filing in the prior application file.

14.(**X**) Please address all further communications to

White & Case LLP
Patent Department
1155 Avenue of the Americas
New York, New York 10036
(212) 819-8200

Respectfully Submitted,

Thelma A. Chen Cleland

Date: October 13, 1999

Thelma A. Chen Cleland
Reg. No. 40,948

White & Case LLP
Patent Department
1155 Ave. of the Americas
New York, NY 10036

(212)819-8200

Enclosures

1103326-0206

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Goutam Das

Serial No.: To be assigned

Filed: Concurrently herewith

Title: DNA MOLECULES FOR EXPRESSION OF POLYPEPTIDES

"Express Mail" Label No. EJ064077753 US.

Date of Deposit October 13, 1999. I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the: Assistant Commissioner for Patents, Washington, D.C. 20231.

Jermaine Peebles
(Type or print name of person mailing paper or fee)

Jermaine Peebles
(Signature of person mailing paper or fee)

Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

Prior to Examination on its merits, please amend the application as follows:

In the Specification:

Please replace pages 22 - 33 containing the paper copy of the Sequence Listing, with enclosed pages 22 - 35. Accordingly, please renumber subsequent pages following the Sequence Listing.

10/13/99 3:36pm
0000BR0C.W51

1103326-0206

REMARKS

The specification has been amended to substitute pages 22-33 containing the paper copy of the Sequence Listing with new enclosed pages 22-35. The amendment only updates the General Information of the Sequence Listing in accordance with 37 C.F.R. § 1.821 - 1.824. A computer readable copy of a Sequence Listing is also enclosed.

Applicant submits that no new matter is presented by the Preliminary Amendment.

In compliance with 37 C.F.R. §§ 1.821 - 1.825, Applicant asserts that the content of the computer readable copy is identical to that of the paper copy of the Sequence Listing submitted herewith.

Applicant requests favorable consideration and entry hereof.

Dated: October 13, 1999

Respectfully submitted,

Thelma A. Chen Cleland

Thelma A. Chen Cleland
Reg. No.: 40,948

White & Case LLP
Patent Department
1155 Avenue of the Americas
New York, New York 10036-2787
(212) 819-8515

Enclosures

10/13/99 3:36pm
0000BR0C.W51

ASTRA AKTIEBOLAG
Södertälje, Sweden

Inventor: DAS, Goutam

20 February 1996

Our ref: HX 1258-1

LH

DNA MOLECULES FOR EXPRESSION OF POLYPEPTIDES

DNA MOLECULES FOR EXPRESSION OF POLYPEPTIDES

TECHNICAL FIELD

The invention relates to DNA molecules, recombinant vectors and cell cultures for use in methods for expression of bile salt-stimulated lipase (BSSL) in the methylotrophic yeast *Pichia pastoris*.

BACKGROUND ART

Bile salt-stimulated lipase (BSSL; EC 3.1.1.1) (for a review see Wang & Hartsuck, 1993) accounts for the majority of the lipolytic activity of the human milk. A characteristic feature of this lipase is that it requires primary bile salts for activity against emulsified long chain triacylglycerols. BSSL has so far been found only in milk from man, gorilla, cat and dog (Hernell et al., 1989).

BSSL has been attributed a critical role for the digestion of milk lipids in the intestine of the breastfed infant (Fredrikzon et al., 1978). BSSL is synthesized in humans in the lactating mammary gland and secretes with milk (Bläckberg et al., 1987). It accounts for approximately 1% of the total milk protein (Bläckberg & Hernell, 1981).

It has been suggested that BSSL is the major rate limiting factor in fat absorption and subsequent growth by, in particular premature, infants who are deficient in their own production of BSSL, and that supplementation of formulas with the purified enzyme significantly improves digestion and growth of these infants (US 4,944,944; Oklahoma Medical Research Foundation). This is clinically important in the preparation of infant formulas which contain relative high percentage of triglycerides and which are based on plant or non human milk protein

sources, since infants fed with these formulas are unable to digest the fat in the absence of added BSSL.

The cDNA structures for both milk BSSL and pancreas carboxylic ester hydrolase (CEH) have been characterized (Baba et al., 1991; Hui and Kissel, 1991; Nilsson et al., 1991; Reue et al., 1991) and the conclusion has been drawn that the milk enzyme and the pancreas enzyme are products of the same gene, the CEL gene. The cDNA sequence (SEQ ID NO: 1) of the CEL gene is disclosed in US 5,200,183 (Oklahoma Medical Research Foundation); WO 91/18293 (Aktiebolaget Astra); Nilsson et al., (1990); and Baba et al., (1991). The deduced amino acid sequence of the BSSL protein, including a signal sequence of 23 amino acids, is shown as SEQ ID NO: 2 in the Sequence Listing, while the sequence of the native protein of 722 amino acids is shown as SEQ ID NO: 3.

The C-terminal region of the protein contains 16 repeats of 11 amino acid residues each, followed by an 11 amino acid conserved stretch. The native protein is highly glycosylated and a large range of observed molecular weights have been reported. This can probably be explained by varying extent of glycosylation (Abouakil et al., 1988). The N-terminal half of the protein is homologous to acetyl choline esterase and some other esterases (Nilsson et al., 1990).

Recombinant BSSL can be produced by expression in a suitable host such as *E. coli*, *Saccharomyces cerevisiae*, or mammalian cell lines. For the scaling-up of a BSSL expression system to make the production cost commercially viable, utilization of heterologous expression systems could be envisaged. As mentioned above, human BSSL has 16 repeats of 11 amino acids at the C-terminal end. To determine the biological significance of this repeat region, various mutants of human BSSL have been constructed which lack part or whole of the repeat regions (Hansson et al., 1993). The variant BSSL-C (SEQ ID NO: 4), for example,

has deletions from amino acid residues 536 to 568 and from amino acid residues 591 to 711. Expression studies, using mammalian cell line C127 host and bovine papilloma virus expression vector, showed that the various variants can be expressed in active forms (Hansson et al., 1993). From the expression studies it was also concluded that the proline rich repeats in human BSSL are not essential for catalytic activity or bile salt activation of BSSL. However, production of BSSL or its mutants in a mammalian expression system could be too expensive for routine therapeutic use.

A eukaryotic system such as yeast may provide significant advantages, compared to the use of prokaryotic systems, for the production of certain polypeptides encoded by recombinant DNA. For example, yeast can generally be grown to higher cell densities than bacteria and may prove capable of glycosylating expressed polypeptides, where such glycosylation is important for the biological activity. However, use of the yeast *Saccharomyces cerevisiae* as a host organism often leads to poor expression levels and poor secretion of the recombinant protein (Cregg et al., 1987). The maximum levels of heterologous proteins in *S. cerevisae* are in the region of 5% of total cell protein (Kingsman et al., 1985). A further drawback of using *Sacharomyces cerevisiae* as a host is that the recombinant proteins tend to be overglycosylated which could affect activity of glycosylated mammalian proteins.

Pichia pastoris is a methylotrophic yeast which can grow on methanol as a sole carbon and energy source as it contains a highly regulated methanol utilization pathway (Ellis et al., 1985). *P. pastoris* is also amenable to efficient high cell density fermentation technology. Therefore recombinant DNA technology and efficient methods of yeast transformation have made it possible to develop *P. pastoris* as a host for expression of heterologous protein in large quantity, with a methanol oxidase promoter based expression system (Cregg et al., 1987).

Use of *Pichia pastoris* is known in the art as a host for the expression of e.g. the following heterologous proteins: human tumor necrosis factor (EP-A-0263311); *Bordetella* pertactin antigens (WO 91/15571); hepatitis B surface antigen (Cregg et al., 1987); human lysozyme protein (WO 92/04441); aprotinin (WO 92/01048). However, successful expression of a heterologous protein in active, soluble and secreted form depends on a variety of factors, e.g. correct choice of signal peptide, proper construction of the fusion junction between the signal peptide and the mature protein, growth conditions, etc.

PURPOSE OF THE INVENTION

The purpose of the invention is to overcome the above mentioned drawbacks with the previous systems and to provide a method for the production of human BSSL with is cost-effective and has a yield comparable with, or superior to, production in other organisms. This purpose has been achieved by providing methods for expression of BSSL in *Pichia pastoris* cells.

By the invention it has thus been shown that human BSSL and the variant BSSL-C can be expressed in active form secreted from *P. pastoris*. The native signal peptide, as well as the heterologous signal peptide derived from *S. cerevisiae* invertase protein, have been used to translocate the mature protein into the culture medium as an active, properly processed form.

DESCRIPTION OF THE INVENTION

In a first aspect, the invention provides a DNA molecule comprising:
(a) a region coding for a polypeptide which is human BSSL or a biologically active variant thereof;
(b) joined to the 5′-end of said polypeptide coding region, a region coding for a signal peptide capable of directing secretion of said polypeptide from *Pichia pastoris* cells transformed with said DNA molecule; and
(c) operably-linked to said coding regions defined in (a) and (b), the methanol oxidase promoter of *Pichia pastoris* or a functionally equivalent promoter.

The term "biologically active variant" of BSSL is to be understood as a polypeptide having BSSL activity and comprising part of the amino acid sequence shown as SEQ ID NO: 3 in the Sequence Listing. The term "polypeptide having BSSL activity" is in this context to be understood as a polypeptide comprising the following properties: (a) being suitable for oral administration; (b) being activated by specific bile-salts; and (c) acting as a non-specific lipase in the contents of the small intestines, i.e. being able to hydrolyze lipids relatively independent of their chemical structure and physical state (emulsified, micellar, soluble).

The said BSSL variant can e.g. be a variant which comprises less than 16 repeat units, whereby a "repeat unit" will be understood as a repeated unit of 11 amino acids, encoded by a nucleotide sequence indicated as a "repeat unit" under the heading "(ix) FEATURE" in "INFORMATION FOR SEQ ID NO: 1" in the Sequence Listing. In particular, the BSSL variant can be the variant BSSL-C, wherein amino acids 536 to 568 and 591 to 711 have been deleted (SEQ ID NO: 4 in the Sequence Listing).

Consequently, the DNA molecule according to the invention is preferably a DNA molecule which encodes BSSL (SEQ ID NO: 3) or BSSL-C (SEQ ID NO: 4).

However, the DNA molecules according to the invention are not to be limited strictly to DNA molecules which encode polypeptides with amino acid sequences identical to SEQ ID NO: 3 or 4 in the Sequence Listing. Rather the invention encompasses DNA molecules which code for polypeptides carrying modifications like substitutions, small deletions, insertions or inversions, which polypeptides nevertheless have substantially the biological activities of BSSL. Included in the invention are consequently DNA molecules coding for BSSL variants as stated above and also DNA molecules coding for polypeptides, the amino acid sequence of which is at least 90% homologous, preferably at least 95% homologous, with the amino acid sequence shown as SEQ ID NO: 3 or 4 in the Sequence Listing.

The signal peptide referred to above can be a peptide which is identical to, or substantially similar to, the peptide with the amino acid sequence shown as amino acids −20 to −1 of SEQ ID NO: 2 in the Sequence Listing. Alternatively, it can be a peptide which comprises a *Saccharomyces cerevisiae* invertase signal peptide.

In a further aspect, the invention provides a vector comprising a DNA molecule as defined above. Preferably, such a vector is a replicable expression vector which carries and is capable of mediating expression, in a cell of the genus *Pichia*, of a DNA sequence coding for human BSSL or a biologically active variant thereof. Such a vector can e.g. be the plasmid vector pARC 5771 (NCIMB 40721), pARC 5799 (NCIMB 40723) or pARC 5797 (NCIMB 40722).

In another aspect, the invention provides a host cell culture comprising cells of the genus *Pichia* transformed with a DNA molecule or a vector as defined above. Preferably, the host cells are *Pichia pastoris* cells of a strain such as PPF-1 or GS115. The said cell culture can e.g. be the culture PPF-1[pARC 5771] (NCIMB 40721), GS115[pARC 5799] (NCIMB 40723) or GS115[pARC 5797] (NCIMB 40722).

In yet another aspect, the invention provides a process the production of a polypeptide which is human BSSL, or a biologically active variant thereof, which comprises culturing host cells according to tne invention under conditions whereby said polypeptide is secreted into the culture medium, and recovering said polypeptide from the culture medium.

EXAMPLES OF THE INVENTION

EXAMPLE 1: Expression of BSSL in *Pichia pastoris* PPF-1

1.1. Construction of pARC 0770

The cDNA sequence (SEQ ID NO: 1) coding for the BSSL protein, including the native signal peptide (below referred to as NSP) was cloned in pTZ19R (Pharmacia) as an *Eco*RI-*Sac*I fragment. The cloning of NSP-BSSL cDNA into *S. cerevisiae* expression vector pSCW 231 (obtained from professor L. Prakash, University of Rochester, NY, USA), which is a low copy number yeast expression vector wherein expression is under control of the constitutive ADH1 promoter, was achieved in two steps. Initially the NSP-BSSL cDNA was cloned into pYES 2.0 (Invitrogen, USA) as an *Eco*RI-*Sph*I fragment from pTZ19R-SP-BSSL. The excess 89 base pairs between the *Eco*RI and *Nco*I at the beginning of the signal peptide coding sequence were removed by creating an *Eco*RI/*Nco*I (89) fusion and regenerating an *Eco*RI site. The resulting clone pARC 0770

contained an ATG codon, originally encoded within the *Nco*I site which was immediately followed by the regenerated *Eco*RI site in frame with the remaining NSP-BSSL sequence.

1.2. Construction of pARC 5771 plasmid

To construct a suitable expression vector for the expression of BSSL, the cDNA fragment encoding the BSSL protein along with its native signal peptide was cloned with *P. pastoris* expression vector pDM 148. The vector pDM 148 (received from Dr. S. Subramani, UCSD) was constructed as follows: the upstream untranslated region (5′-UTR) and the down stream untranslated region (3′-UTR) of methanol oxidase (MOX1) gene were isolated by PCR and placed in tandem in the multiple cloning sequence (MCS) of *E. coli* vector pSK$^+$ (available from Stratagene, USA).

For proper selection of the putative *P. pastoris* transformants, a DNA sequence coding for *S. cerevisiae* ARG4 gene along with its own promoter sequence was inserted between the 5′- and the 3′-UTR in pSK-. The resulting construct pDM148 has following features: in the MCS region of pSK- the 5′-UTR of MOX, *S. cerevisiae* ARG4 genomic sequence and the 3′-UTR of MOX were cloned. Between the 5′-UTR of MOX and the ARG4 genomic sequence a series of unique restriction sites (*Sal*I, *Cla*I, *Eco*RI, *Pst*I, *Sma*I and *Bam*HI) were situated where any heterologous protein coding sequence can be cloned for expression under the control of the MOX promoter in *P. pastoris*. To facilitate integration of this expression cassette into the MOX1 locus in *P. pastoris* chromosome, the expression cassette can be cleaved from the rest of the pSK$^-$ vector by digestion with *Not*I restriction enzyme.

The 5′-UTR of MOX1 of *P. pastoris* cloned in pDM 148 was about 500 bp in length while the 3′-UTR of MOX1 from *P. pastoris* cloned into pDM

148 was about 1000 bp long. To insert the NSP-BSSL cDNA sequence, between the 5'-UTR of MOX1 and the *S. cerevisiae* ARG4 coding sequence in pDM 148, the cDNA insert (SP-BSSL) was isolated from pARC 0770 by digestion with *Eco*RI and *Bam*HI (approximately 2.2 kb DNA fragment) and cloned between the *Eco*RI and *Bam*HI sites in pDM 148.

The resulting construct pARC 5771 (NCIMB 40721) contained the *P. pastoris* MOX1 5'-UTR followed by the NSP-BSSL coding sequence followed by *S. cerevisiae* ARG4 gene sequence and 3'-UTR of MOX1 gene of *P. pastoris* while the entire DNA segment from 5'-UTR of MOX1 to the 3'-UTR of MOX1 was cloned at the MCS of pSK-.

1.3. Transformation of BSSL in *P. pastoris* host PPF-1

For expression of BSSL in *P. pastoris* host PPF-1 (his4, arg4; received from Phillips Petroleum Co.), the plasmid pARC 5771 was digested with *Not*I and the entire digested mix (10 μg of total DNA) was used to transform PPF-1. The transformation protocol followed was essentially the yeast spheroplast method described by Cregg et al. (1987). Transformants were regenerated on minimal medium lacking arginine so that Arg+ colonies could be selected. The regeneration top agar containing the transformants was lifted and homogenized in water and yeast cells plated to about 250 colonies per plate on minimal glucose plates lacking arginine. Mutant colonies are then identified by replica plating onto minimal methanol plates. Approximately 15% of all transformants turned out to be Mut^S (methanol slow growing) phenotype.

1.4. Screening for transformants expressing BSSL

In order to screen large number of transformants rapidly for the expression of lipase a lipase plate assay method was developed. The procedure for preparing these plates was as follows: to a solution of 2% agarose (final), 10 x Na-cholate solution in water was added to a final concentration of 1%. The lipid substrate trybutine was added in the mixture to a final concentration of 1% (v/v). To support growth of the transformants the mixture was further supplemented with 0.25% yeast nitrogen base (final) and 0.5% methanol (final). The ingredients were mixed properly and poured into plates upto 3-5 mm thickness. Once the mixture became solid, the transformants were streaked onto the plates and the plates were further incubated at +37°C for 12 h. The lipase producing clones showed a clear halo around the clone. In a typical experiment 7 out of a total of 93 transformants were identified as BSSL producing transformants. Two clones (Nos. 39 and 86) producing the largest halos around the streaked colony were picked out for further characterization.

1.5. Expression of BSSL from PPF-1[pARC 5771]

The two transformants Nos. 39 and 86 described in Section 1.4 were picked out and grown in BMGY liquid media (1% yeast extract, 2% bactopeptone, 1.34% yeast nitrogen base without amino acid, 100 mM KPO_4 buffer, pH 6.0, 400 µg/l biotin, and 2% glycerol) for 24 h at 30°C until the cultures reached A_{600} close to 40. The cultures were pelleted down and resuspended in BMMY (2% glycerol replaced by 0.5% methanol in BMGY) media at A_{600} = 300. The induced cultures were incubated at 30°C with shaking for 120 h. The culture supernatants were withdrawn at different time points for the analysis of the expression of BSSL by enzyme activity assay, SDS-PAGE analysis and western blotting.

1.6. Detection of BSSL enzyme activity in the culture supernatants of clone Nos. 39 and 86

To determine the enzyme activity in the cell free culture supernatant of the induced cultures Nos. 39 and 86 as described in Section 1.5, the cultures were spun down and 2 µl of the cell free supernatant was assayed for BSSL enzyme activity according to the method described by Hernell and Olivecrona (1974). As shown in Table 1, both the cultures were found to contain BSSL enzyme activity with the maximum activity at 96 h following induction.

1.7. Western blot analysis of culture supernatants of PPF-1:pARC 5771 transformants (Nos. 39 and 86)

To determine the presence of recombinant BSSL in the culture supernatants Nos. 39 and 86 of PPF-1[pARC 5771] transformants, the cultures were grown and induced as described in Section 1.5. The cultures were withdrawn at different time points following induction and subjected to Western blot analysis using anti BSSL polyclonal antibody. The results indicated the presence of BSSL in the culture supernatant as a 116 kDa band.

EXAMPLE 2: Expression of BSSL in *Pichia pastoris* GS115

2.1. Construction of pARC 5799

Since the 5′-MOX UTR and 3′-MOX UTR were not properly defined and since the pDM 148 vector lacks any other suitable marker (e.g. a G418 resistance gene) to monitor the number of copies of the BSSL integrated in the *Pichia* chromosome, the cDNA insert of native BSSL along with its signal peptide was cloned into another *P. pastoris* expression vector, pHIL D4. The integrative plasmid pHIL D4 was obtained from Phillips

Petroleum Company. The plasmid contained 5'-MOX1, approximately 1000 bp segment of the alcohol oxidase promoter and a unique *Eco*RI cloning site. It also contained approximately 250 bp of 3'-MOX1 region containing alcohol oxidase terminating sequence, following the *Eco*RI site. The "termination" region was followed by *P. pastoris* histidinol dehydrogenase gene *HIS4* contained on a 2.8 kb fragment to complement the defective *HIS4* gene in the host GS115 (see below). A 650 bp region containing 3'-MOX1 DNA was fused at the 3'-end of HIS4 gene, which together with the 5'-MOX1 region was necessary for site-directed integration. A bacterial kanamycin resistance gene from pUC-4K (PL-Biochemicals) was inserted at the unique *Nae*I site between *HIS4* and 3'-MOX1 region at 3' of the *HIS4* gene.

To clone the NSP-BSSL coding cDNA fragment at the unique *Eco*RI site of pHIL D4, a double stranded oligo linker having a *Bam*HI–*Eco*RI cleaved position was ligated to the *Bam*HI digested plasmid pARC 5771 and the entire NSP-BSSL coding sequence was pulled out as a 2.2 kb *Eco*RI fragment. This fragment was cloned at the *Eco*RI site of pHIL D-4 and the correctly oriented plasmid was designated as pARC 5799 (NCIMB 40723).

2.2. Transformation of pARC 5799

To facilitate integration of the NSP-BSSL coding sequence at the genomic locus of MOX1 in *P. pastoris* the plasmid pARC 5799 was digested with *Bgl*II and used for transformation of *P. pastoris* strain GS115(his4) (Phillips Petroleum Company) according to a protocol described in Section 1.5. In this case, however, the selection was for His prototrophy. The transformants were picked up following serial dilution plating of the regenerated top agar and tested directly for lipase plate assay as described in Section 1.4. Two transformant clones (Nos. 9 and 21) were picked up on the basis of the halo size on the lipase assay plate and

checked further for the expression of BSSL. The clones were found to be Mut^+.

2.3. Determination of BSSL enzyme activity in the culture supernatants of GS115[pARC 5799] transformants Nos. 9 and 21.

The two transformed clones Nos. 9 and 21 of GS115[pARC 5799] were grown essentially following the protocol described in Section 1.5. The culture supernatants at different time points following induction were assayed for BSSL enzyme activity as described in Section 1.6. As shown in Table 1, both the culture supernatants were found to contain BSSL enzyme activity and the enzyme activity was highest after 72 h of induction. Both clones showed a superior expression of BSSL compared to the clones of PPF-1[pARC 5771].

2.4. SDS-PAGE and western blot analysis of culture supernatants of GS115[pARC 5799] transformants Nos. 9 and 21

The culture supernatants collected at different time points, as described in Section 2.3 were subjected to SDS-PAGE and western blot analysis. From the SDS-PAGE profile it was estimated that about 60-75% of the total protein present in the culture supernatants of the induced cultures was BSSL. The molecular weight of the protein was about 116 kDa. The western blot data also confirmed that the major protein present in the culture supernatant was BSSL. The protein apparently had the same molecular weight as the native BSSL.

EXAMPLE 3: Scaling-up of BSSL expression

3.1. Scaling-up of expression of BSSL from the transformed clone GS115[pARC 5799] (No. 21)

A 23 l capacity B. Braun fermenter was used. Five litres of medium containing, 1% YE, 2% Peptone, 1.34 YNB and 4% w/v glycerol was autoclaved at 121°C for 30 min and biotin (400 µg/L final concentration) was added during inoculation after filter sterilization. For inoculum, glycerol stock of GS115[pARC 5799] (No. 21) inoculated into a synthetic medium containing YNB (67%) plus 2% glycerol (150 ml) and grown at +30°C for 36 h was used. Fermentation conditions were as follows: the temperature was +30°C; pH 5.0 was maintained using 3.5 N NH_4OH and 2 N HCl; dissolved oxygen from 20 to 40% of air saturation; polypropylene glycol 2000 was used as antifoam.

Growth was monitored at regular intervals by taking OD at 600 nm. A_{600} reached a maximum of 50-60 in 24 h. At this point, the batch growth phase was over as indicated by the increased dissolved oxygen levels.

Growth phase was immediately followed by the induction phase. During this phase, methanol containing 12 ml/L PTM1 salts was fed. Methanol feed rate was 6 µl/h during first 10-12 h after which it was increased gradually in 6 ml/h increments every 7-8 h to a maximum of 36 ml/h. Ammonia used for pH control acted as a nitrogen source. Methanol accumulation was checked every 6-8 h by using dissolved oxygen spiking and it was found to be limiting during the entire phase of induction. OD at 600 nm increased from 50-60 to 150-170 during 86 h of methanol feed. Yeast extract and peptone were added every 24 h to make final conc. of 0.25% and 0.5% respectively.

Samples were withdrawn at 24 h interval and checked for BSSL enzyme activity in the cell free broth. The broth was also subjected to SDS-PAGE and western blotting analysis.

3.2. Protein analysis of the secreted BSSL from the fermenter grown culture GS115[pARC 5799] (No. 21)

BSSL enzyme activity in cell free broth increased from 40-70 mg/l (equivalent of native protein) in 24 h to a maximum 200-227.0 mg/l (equivalent of native protein) at the end of 86-90 h. SDS-PAGE analysis of the cell free broth shows a prominent coomassie blue stained band of mol.wt. of 116 kDa. The identity of the band was confirmed by Western blot performed as described in Section 1.7 for native BSSL.

3.3. Purification of recombinant BSSL secreted into the culture supernatant of GS115[pARC 5799] (No. 21) clones

The *P. pastoris* clone GS115[pARC 5799] was grown and induced in the fermenter as described in Section 3.1. For purification of recombinant BSSL, 250 ml of culture medium (induced for 90 h) was spun at 12,000 x g for 30 minutes to remove all particulate matter. The cell free culture supernatant was ultra filtered in an Amicon set up using a 10 kDa cut off membrane. Salts and low molecular weight proteins and peptides of the culture supernatant were removed by repeated dilution during filtration. The buffer used for such dilution was 5 mM Barbitol pH 7.4. Following concentration of the culture supernatant, the retentate was reconstituted to 250 ml using 5 mM Barbitol, pH 7.4 and 50 mM NaCl and loaded onto a Heparin-Sepharose column (15 ml bed volume) which was pre-equilibrated with the same buffer. The sample loading was done at a flow rate of 10 ml/hr. Following loading the column was washed with 5 mM Barbitol, pH 7.4 and 0.1 M NaCl (200 µl washing buffer) till the absorbance at 250 nm reached below detection level. The BSSL was eluted with 200 ml of Barbitol buffer (5 mM, pH 7.4) and a linear gradient of NaCl ranging from 0.1 M to 0.7 M. Fractions (2.5 ml) were collected and checked for the eluted protein by monitoring the absorbance at 260 nm. Fractions containing protein were assayed for

BSSL enzyme activity. Appropriate fractions were analyzed on 8.0% SDS-PAGE to check thee purification profile.

3.4. Characterization of purified recombinant BSSL secreted in the culture supernatant of GS115[pARC 5799]

SDS-PAGE and Western blot analysis of the fractions (described in Section 3.3) showing maximal BSSL enzyme activity demonstrated that the recombinant protein was approximately 90% pure. The molecular weight of the purified protein was about 116 κDa as determined by SDS-PAGE and western blot analysis. When the samples were overloaded for SDS-PAGE analysis a low molecular weight protein band could be detected by Coomassie Brilliant Blue staining which was not picked up on Western blot. The purified protein was subjected to N-terminal analysis in an automated protein sequencer. The results showed that the protein was properly processed from the native signal peptide and the recombinant protein has the N-terminal sequence A K L G A V Y. The specific activity of the purified recombinant protein was found to be similar to that of the native protein.

EXAMPLE 4: Expression of BSSL-C in *Pichia pastoris* GS115

4.1. Construction of pARC 5797

The cDNA coding sequence for the BSSL variant BSSL-C was fused at its 5′-end with the signal peptide coding sequence of *S. cerevisiae* SUC2 gene product (invertase), maintaining the integrity of the open reading frame initiated at the first ATG codon of invertase signal peptide. This fusion gene construct was initially cloned into the *S. cerevisiae* expression vector pSCW 231 (pSCW 231 is a low copy number yeast expression vector and the expression is under the control of the constitutive ADH1

promoter) between *Eco*RI and *Bam*HI site to generate the expression vector pARC 0788.

The cDNA of the fusion gene was further subcloned into *P. pastoris* expression vector pDM 148 (described in Section 1.2) by releasing the appropriate 1.8 kb fragment by *Eco*RI and *Bam*HI digestion of pARC 0788 and subcloning the fragment into pDM 148 digested with *Eco*RI and *Bam*HI. The resulting construct pARC 5790 was digested with *Bam*HI and a double stranded oligonucleotide linker of the physical structure *Bam*HI–*Eco*RI–*Bam*HI was ligated to generate the construct pARC 5796 essentially to isolate the cDNA fragment of the fusion gene, following the strategy as described in Section 2.1.

Finally the 1.8 kb fragment containing the invertase signal peptide / BSSL-C fusion gene was released from pARC 5796 by *Eco*RI digestion and cloned into pHIL D4 at the *Eco*RI site. By appropriate restriction analysis of the expression vector containing the insert in the proper orientation was identified and was designated as pARC 5797 (NCIMB 40722).

4.2. Expression of recombinant BSSL-C from *P. pastoris*

To express recombinant BSSL-C from *P. pastoris*, the *P. pastoris* host GS115 was transformed with pARC 5797 by the method as described in Sections 1.3 and 2.2. Transformants were checked for lipase production by the method described in Sections 1.4 and 2.2. A single transformant (No. 3) was picked on the basis of high lipase producing ability by the lipase plate assay detection method and was further analyzed for production of BSSL enzyme activity in the culture supernatant by essentially following the method as described in Sections 1.6 and 2.3. As shown in Table 1, the culture supernatant of GS115[pARC 5797] (No. 3)

contained BSSL enzyme activity and the amount increased progressively till 72 h following induction.

4.3. SDS-PAGE and western blot analysis of culture supernatant of GS115[pARC 5797] transformant (No. 3)

The culture supernatant collected at various time points as described in Section 4.2 were subjected to SDS-PAGE and western blot analysis as described in Sections 1.7 and 2.4. From the SDS-PAGE profile it was estimated that about 75-80% of the total extracellular protein was BSSL-C. The molecular weight of the protein as estimated from SDS-PAGE analysis was approximately 66 kDa. On western blot analysis only two bands (doublet) around 66 kDa were found to be immunoreactive and thus confirming the expression of recombinant BSSL-C.

EXAMPLE FOR COMPARISON: Expression of BSSL in *S. cerevisiae*

Attempts to express BSSL in *Saccharomyces cerevisiae* were made. BSSL was poorly secreted in *S. cerevisiae* and the native signal peptide did not work efficiently. In addition, the native signal peptide did not get cleaved from the mature protein in *S. cerevisiae*.

REFERENCES

Abouakil, N., Rogalska, E., Bonicel, J. and Lombardo, D. (1988) Biochim. Biophys. Acta. 961, 299-308.

Baba, T., Downs, D., Jackson, K.W., Tang, J. and Wang, C-S (1991) Biochemistry 30, 500-510.

Bläckberg, L. and Hernell, O. (1981) Eur. J. Biochem. 116, 221-225.

Bläckberg, L., Ängquist, K.A. and Hernell, O. (1987) FEBS Lett. 217, 37-41.

Cregg, J.M. et al. (1987) Bio/Technology 5, 479-485.

Ellis, S.B. et al. (1985) Mol. Cell. Biol. 5, 1111-1121.

Fredrikzon, B., Hernell, O., Bläckberg, L. and Olivecrona, T. (1978) Pediatric Res. 12, 1048-1052.

Hansson, L., Bläckberg, L., Edlund, M., Lundberg, L., Strömqvist, M. and Hernell, O. (1993) J. Biol. Chem. 268, 26692-26698.

Hernell, O. and Olivecrona, T. (1974) Biochim. Biophys. Acta 369, 234-244.

Hernell, O., Bläckberg, L and Olivecrona, T. (1989) in: Textbook of gastroenterology and nutrition in infancy (Lebenthal, E., ed.) 347-354, Raven Press, NY.

Hernell, O. and Bläckberg, L. (1982) Pediatric Res. 16, 882-885.

Hui, D. Y. and Kissel, J. A. (1990) FEBS Letters 276, 131-134.

Kingsman, et.al. (1985) Biotechnology and Genetic Engineering Reviews 3, 377-416.

Nilsson, J., Bläckberg, L., Carlsson, P., Enerbäck, S., Hernell, O. and Bjursell, G. (1990) Eur. J. Biochem. 192, 543-550.

Reue, K., Zambaux, J., Wong, H., Lee, G., Leete, T.H., Ronk, M., Shively, J.E., Sternby, B., Borgström, B., Ameis, D. and Scholtz, M.C. (1991) J. Lipid. Res. 32, 267-276.

Wang, C-S, and Hartsuck, J.A. (1993) Biochim. Biphys Acta 1166, 1-19.

DEPOSIT OF MICROORGANISMS

The following plasmids, transformed into *Pichia pastoris* cultures, have been deposited under the Budapest Treaty at the National Collections of Industrial and Marine Bacteria (NCIMB), Aberdeen, Scotland, UK. The date of deposit is 2 May 1995.

Strain[plasmid]	NCIMB No.
PPF-1[pARC 5771]	40721
GS115[pARC 5799]	40723
GS115[pARC 5797]	40722

TABLE 1

Enzyme activity in the culture supernatants of *Pichia pastoris* transformants.

Hours after induction	Enzyme activity in mg/L equivalent of native BSSL				
	PPF-1[pARC 5771]		GS115[pARC 5799]		GS115[pARC 5797]
	No. 39	No. 86	No. 9	No. 21	No. 3
24	0.254	0.135	1.53	1.72	0.37
48	2.69	3.12	17.28	34.70	40.9
72	3.96	8.25	37.37	50.60	44.9
96	11.26	13.60	26.34	50.60	35.6
120	8.42	13.13	13.60	22.30	17.8

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: ASTRA AB
(B) STREET: Västra Mälarehamnen 9
(C) CITY: Södertälje
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP): S-151 85
(G) TELEPHONE: +46-8-553 260 00
(H) TELEFAX: +46-8-553 288 20
(I) TELEX: 19237 astra s

(ii) TITLE OF INVENTION: DNA Sequences for Expression of Polypeptides

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2428 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: mammary gland

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:82..2319
(D) OTHER INFORMATION:/product= "bile-salt-stimulated lipase"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:985..1173

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:1174..1377

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:1378..1575

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:1576..2415

(ix) FEATURE:
    (A) NAME/KEY: mat_peptide
    (B) LOCATION:151..2316

(ix) FEATURE:
    (A) NAME/KEY: polyA_signal
    (B) LOCATION:2397..2402

(ix) FEATURE:
    (A) NAME/KEY: repeat_region
    (B) LOCATION:1756..2283

(ix) FEATURE:
    (A) NAME/KEY: 5'UTR
    (B) LOCATION:1..81

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1756..1788

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1789..1821

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1822..1854

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1855..1887

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1888..1920

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1921..1953

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1954..1986

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:1987..2019

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:2020..2052

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:2053..2085

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:2086..2118

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:2119..2151

(ix) FEATURE:
    (A) NAME/KEY: repeat_unit
    (B) LOCATION:2152..2184

```
     (ix) FEATURE:
          (A) NAME/KEY: repeat_unit
          (B) LOCATION:2185..2217

     (ix) FEATURE:
          (A) NAME/KEY: repeat_unit
          (B) LOCATION:2218..2250

     (ix) FEATURE:
          (A) NAME/KEY: repeat_unit
          (B) LOCATION:2251..2283

      (x) PUBLICATION INFORMATION:
          (A) AUTHORS: Nilsson, Jeanette
                       Bläckberg, Lars
                       Carlsson, Peter
                       Enerbäck, Sven
                       Hernell, Olle
                       Bjursell, Gunnar
          (B) TITLE: cDNA cloning of human-milk
                bile-salt-stimulated lipase and evidence for its
                identity to pancreatic carboxylic ester hydrolase
          (C) JOURNAL: Eur. J. Biochem.
          (D) VOLUME: 192
          (F) PAGES: 543-550
          (G) DATE: Sept.-1990

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ACCTTCTGTA TCAGTTAAGT GTCAAGATGG AAGGAACAGC AGTCTCAAGA TAATGCAAAG        60

AGTTTATTCA TCCAGAGGCT G ATG CTC ACC ATG GGG CGC CTG CAA CTG GTT         111
                        Met Leu Thr Met Gly Arg Leu Gln Leu Val
                        -23         -20                 -15

GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC GCG AAG CTG         159
Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala Ala Lys Leu
            -10                 -5                   1

GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA GGC GTC AAT AAG AAG         207
Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val Asn Lys Lys
      5                  10                  15

CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC AAG GGC ATC CCC TTC         255
Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly Ile Pro Phe
 20                  25                  30                  35

GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG CCA CAT CCT GGC TGG         303
Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His Pro Gly Trp
                 40                  45                  50

CAA GGG ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC         351
Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala
             55                  60                  65

ACC ATC ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC         399
Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu
         70                  75                  80

AAC ATT TGG GTG CCC CAG GGC AGG AAG CAA GTC TCC CGG GAC CTG CCC         447
Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg Asp Leu Pro
     85                  90                  95

GTT ATG ATC TGG ATC TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT         495
Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His
100                 105                 110                 115
```

```
GGG GCC AAC TTC CTC AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC      543
Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala
                120                 125                 130

ACA CGC GGA AAC GTC ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC      591
Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro
            135                 140                 145

CTT GGG TTC CTC AGC ACT GGG GAC GCC AAT CTG CCA GGT AAC TAT GGC      639
Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly
        150                 155                 160

CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG AAT ATC GCG      687
Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg Asn Ile Ala
    165                 170                 175

GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG GAG TCT GCT      735
Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala
180                 185                 190                 195

GGA GGT GCC AGC GTC TCT CTG CAG ACC CTC TCC CCC TAC AAC AAG GGC      783
Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly
                200                 205                 210

CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT CCC TGG      831
Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser Pro Trp
            215                 220                 225

GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG GTG GCT GAG AAG      879
Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val Ala Glu Lys
        230                 235                 240

GTG GGT TGC CCT GTG GGT GAT GCC GCC AGG ATG GCC CAG TGT CTG AAG      927
Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln Cys Leu Lys
    245                 250                 255

GTT ACT GAT CCC CGA GCC CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA      975
Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala
260                 265                 270                 275

GGC CTG GAG TAC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT GTC ATT     1023
Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro Val Ile
                280                 285                 290

GAT GGA GAC TTC ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC AAC GCC     1071
Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala
            295                 300                 305

GCC GAC ATC GAC TAT ATA GCA GGC ACC AAC AAC ATG GAC GGC CAC ATC     1119
Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp Gly His Ile
        310                 315                 320

TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC AAC AAG AAA GTC     1167
Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn Lys Lys Val
    325                 330                 335

ACG GAG GAG GAC TTC TAC AAG CTG GTC AGT GAG TTC ACA ATC ACC AAG     1215
Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr Ile Thr Lys
340                 345                 350                 355

GGG CTC AGA GGC GCC AAG ACG ACC TTT GAT GTC TAC ACC GAG TCC TGG     1263
Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp
                360                 365                 370

GCC CAG GAC CCA TCC CAG GAG AAT AAG AAG AAG ACT GTG GTG GAC TTT     1311
Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val Val Asp Phe
            375                 380                 385
```

```
GAG ACC GAT GTC CTC TTC CTG GTG CCC ACC GAG ATT GCC CTA GCC CAG        1359
Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala Leu Ala Gln
        390                 395                 400

CAC AGA GCC AAT GCC AAG AGT GCC AAG ACC TAC GCC TAC CTG TTT TCC        1407
His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser
    405                 410                 415

CAT CCC TCT CGG ATG CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT        1455
His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His
420                 425                 430                 435

GCA GAT GAC ATT CAG TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG        1503
Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr
                440                 445                 450

GGC TAC CGG CCC CAA GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC        1551
Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr
            455                 460                 465

TGG ACC AAC TTT GCC AAA ACA GGG GAC CCC AAC ATG GGC GAC TCG GCT        1599
Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly Asp Ser Ala
        470                 475                 480

GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG GAA AAC AGC GGC TAC CTG        1647
Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser Gly Tyr Leu
    485                 490                 495

GAG ATC ACC AAG AAG ATG GGC AGC AGC TCC ATG AAG CGG AGC CTG AGA        1695
Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg Ser Leu Arg
500                 505                 510                 515

ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC TAT CTG GCG CTG CCC ACA        1743
Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr
                520                 525                 530

GTG ACC GAC CAG GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC GAG        1791
Val Thr Asp Gln Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu
            535                 540                 545

GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC GAG ACC GCC CCC GTG CCG        1839
Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro
        550                 555                 560

CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC        1887
Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser
    565                 570                 575

GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG        1935
Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
580                 585                 590                 595

CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC        1983
Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp
                600                 605                 610

TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC        2031
Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro
            615                 620                 625

GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT        2079
Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
        630                 635                 640

GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC        2127
Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro
    645                 650                 655
```

CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG ACC CCC ACG 2175
Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Thr Pro Thr
660 665 670 675

GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC 2223
Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala
680 685 690

CCC CCT GTG CCC CCC ACG GGT GAC TCT GAG GCT GCC CCT GTG CCC CCC 2271
Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Ala Pro Val Pro Pro
695 700 705

ACA GAT GAC TCC AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT TAG 2319
Thr Asp Asp Ser Lys Glu Ala Gln Met Pro Ala Val Ile Arg Phe *
710 715 720

CGTCCCATGA GCCTTGGTAT CAAGAGGCCA CAAGAGTGGG ACCCCAGGGG CTCCCCTCCC 2379

ATCTTGAGCT CTTCCTGAAT AAAGCCTCAT ACCCCTAAAA AAAAAAAAA 2428

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 746 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Leu Thr Met Gly Arg Leu Gln Leu Val Val Leu Gly Leu Thr Cys
-23 -20 -15 -10

Cys Trp Ala Val Ala Ser Ala Ala Lys Leu Gly Ala Val Tyr Thr Glu
-5 1 5

Gly Gly Phe Val Glu Gly Val Asn Lys Lys Leu Gly Leu Leu Gly Asp
10 15 20 25

Ser Val Asp Ile Phe Lys Gly Ile Pro Phe Ala Ala Pro Thr Lys Ala
30 35 40

Leu Glu Asn Pro Gln Pro His Pro Gly Trp Gln Gly Thr Leu Lys Ala
45 50 55

Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala Thr Ile Thr Gln Asp Ser
60 65 70

Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu Asn Ile Trp Val Pro Gln
75 80 85

Gly Arg Lys Gln Val Ser Arg Asp Leu Pro Val Met Ile Trp Ile Tyr
90 95 100 105

Gly Gly Ala Phe Leu Met Gly Ser Gly His Gly Ala Asn Phe Leu Asn
110 115 120

Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn Val Ile
125 130 135

Val Val Thr Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser Thr
140 145 150

Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly Leu Arg Asp Gln His Met
155 160 165

Ala Ile Ala Trp Val Lys Arg Asn Ile Ala Ala Phe Gly Gly Asp Pro
170 175 180 185

Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala Gly Gly Ala Ser Val Ser
190 195 200

Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly Leu Ile Arg Arg Ala Ile
205 210 215

Ser Gln Ser Gly Val Ala Leu Ser Pro Trp Val Ile Gln Lys Asn Pro
220 225 230

Leu Phe Trp Ala Lys Lys Val Ala Glu Lys Val Gly Cys Pro Val Gly
235 240 245

Asp Ala Ala Arg Met Ala Gln Cys Leu Lys Val Thr Asp Pro Arg Ala
250 255 260 265

Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala Gly Leu Glu Tyr Pro Met
270 275 280

Leu His Tyr Val Gly Phe Val Pro Val Ile Asp Gly Asp Phe Ile Pro
285 290 295

Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala Ala Asp Ile Asp Tyr Ile
300 305 310

Ala Gly Thr Asn Asn Met Asp Gly His Ile Phe Ala Ser Ile Asp Met
315 320 325

Pro Ala Ile Asn Lys Gly Asn Lys Lys Val Thr Glu Glu Asp Phe Tyr
330 335 340 345

Lys Leu Val Ser Glu Phe Thr Ile Thr Lys Gly Leu Arg Gly Ala Lys
350 355 360

Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro Ser Gln
365 370 375

Glu Asn Lys Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe
380 385 390

Leu Val Pro Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys
395 400 405

Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser His Pro Ser Arg Met Pro
410 415 420 425

Val Tyr Pro Lys Trp Val Gly Ala Asp His Ala Asp Asp Ile Gln Tyr
430 435 440

Val Phe Gly Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro Gln Asp
445 450 455

Arg Thr Val Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala Lys
460 465 470

Thr Gly Asp Pro Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu
475 480 485

Pro Tyr Thr Thr Glu Asn Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met
490 495 500 505

Gly Ser Ser Ser Met Lys Arg Ser Leu Arg Thr Asn Phe Leu Arg Tyr
510 515 520

Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln Glu Ala
525 530 535

Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro
540 545 550

Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly
555 560 565

Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro
570 575 580 585

Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser
590 595 600

Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
605 610 615

Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp
620 625 630

Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ala Gly Pro Pro Pro
635 640 645

Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
650 655 660 665

Asp Ser Gly Ala Pro Pro Val Thr Pro Thr Gly Asp Ser Glu Thr Ala
670 675 680

Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr
685 690 695

Gly Asp Ser Glu Ala Ala Pro Val Pro Pro Thr Asp Asp Ser Lys Glu
700 705 710

Ala Gln Met Pro Ala Val Ile Arg Phe *
715 720

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 722 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: Mammary gland

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala Lys Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val
1 5 10 15

Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly
20 25 30

Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His
35 40 45

Pro Gly Trp Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys
50 55 60

```
Leu Gln Ala Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys
65                  70                  75                  80

Leu Tyr Leu Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg
                85                  90                  95

Asp Leu Pro Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly
            100                 105                 110

Ser Gly His Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu
        115                 120                 125

Glu Ile Ala Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg
    130                 135                 140

Val Gly Pro Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly
145                 150                 155                 160

Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg
                165                 170                 175

Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly
            180                 185                 190

Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr
        195                 200                 205

Asn Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu
    210                 215                 220

Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val
225                 230                 235                 240

Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln
                245                 250                 255

Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val
            260                 265                 270

Pro Leu Ala Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val
        275                 280                 285

Pro Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr
    290                 295                 300

Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp
305                 310                 315                 320

Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn
                325                 330                 335

Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr
            340                 345                 350

Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr
        355                 360                 365

Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val
    370                 375                 380

Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala
385                 390                 395                 400

Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr
                405                 410                 415

Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly
            420                 425                 430
```

```
Ala Asp His Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala
        435                 440                 445

Thr Pro Thr Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met
    450                 455                 460

Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly
465                 470                 475                 480

Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser
                485                 490                 495

Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg
            500                 505                 510

Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala
        515                 520                 525

Leu Pro Thr Val Thr Asp Gln Glu Ala Thr Pro Val Pro Pro Thr Gly
    530                 535                 540

Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Thr Ala
545                 550                 555                 560

Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr
                565                 570                 575

Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala
            580                 585                 590

Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro
        595                 600                 605

Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly
    610                 615                 620

Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro
625                 630                 635                 640

Pro Thr Gly Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp Ser
                645                 650                 655

Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
            660                 665                 670

Thr Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp
        675                 680                 685

Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Ala Pro
    690                 695                 700

Val Pro Pro Thr Asp Asp Ser Lys Glu Ala Gln Met Pro Ala Val Ile
705                 710                 715                 720

Arg Phe
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 568 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Homo sapiens
    (F) TISSUE TYPE: Mammary gland

(ix) FEATURE:
    (A) NAME/KEY: Peptide
    (B) LOCATION:1..568
    (D) OTHER INFORMATION:/label= Variant_C

(x) PUBLICATION INFORMATION:
    (A) AUTHORS: Hansson, Lennart
        Blackberg, Lars
        Edlund, Michael
        Lundberg, Lennart
        Stromqvist, Mats
        Hernell, Olle
    (B) TITLE: Recombinant Human Milk Bile Salt-stimulated Lipase
    (C) JOURNAL: J. Biol. Chem.
    (D) VOLUME: 268
    (E) ISSUE: 35
    (F) PAGES: 26692-26698
    (G) DATE: Dec. 15-1993

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Ala	Lys	Leu	Gly	Ala	Val	Tyr	Thr	Glu	Gly	Gly	Phe	Val	Glu	Gly	Val
1				5					10					15	
Asn	Lys	Lys	Leu	Gly	Leu	Leu	Gly	Asp	Ser	Val	Asp	Ile	Phe	Lys	Gly
			20					25					30		
Ile	Pro	Phe	Ala	Ala	Pro	Thr	Lys	Ala	Leu	Glu	Asn	Pro	Gln	Pro	His
		35					40					45			
Pro	Gly	Trp	Gln	Gly	Thr	Leu	Lys	Ala	Lys	Asn	Phe	Lys	Lys	Arg	Cys
	50					55					60				
Leu	Gln	Ala	Thr	Ile	Thr	Gln	Asp	Ser	Thr	Tyr	Gly	Asp	Glu	Asp	Cys
65					70					75					80
Leu	Tyr	Leu	Asn	Ile	Trp	Val	Pro	Gln	Gly	Arg	Lys	Gln	Val	Ser	Arg
				85					90					95	
Asp	Leu	Pro	Val	Met	Ile	Trp	Ile	Tyr	Gly	Gly	Ala	Phe	Leu	Met	Gly
			100					105					110		
Ser	Gly	His	Gly	Ala	Asn	Phe	Leu	Asn	Asn	Tyr	Leu	Tyr	Asp	Gly	Glu
		115					120					125			
Glu	Ile	Ala	Thr	Arg	Gly	Asn	Val	Ile	Val	Val	Thr	Phe	Asn	Tyr	Arg
	130					135					140				
Val	Gly	Pro	Leu	Gly	Phe	Leu	Ser	Thr	Gly	Asp	Ala	Asn	Leu	Pro	Gly
145					150					155					160
Asn	Tyr	Gly	Leu	Arg	Asp	Gln	His	Met	Ala	Ile	Ala	Trp	Val	Lys	Arg
				165					170					175	
Asn	Ile	Ala	Ala	Phe	Gly	Gly	Asp	Pro	Asn	Asn	Ile	Thr	Leu	Phe	Gly
			180					185					190		
Glu	Ser	Ala	Gly	Gly	Ala	Ser	Val	Ser	Leu	Gln	Thr	Leu	Ser	Pro	Tyr
		195					200					205			

```
Asn Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu
    210                 215                 220

Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val
225                 230                 235                 240

Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln
                245                 250                 255

Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val
            260                 265                 270

Pro Leu Ala Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val
        275                 280                 285

Pro Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr
    290                 295                 300

Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp
305                 310                 315                 320

Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn
                325                 330                 335

Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr
            340                 345                 350

Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr
        355                 360                 365

Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val
    370                 375                 380

Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala
385                 390                 395                 400

Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr
                405                 410                 415

Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly
            420                 425                 430

Ala Asp His Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala
        435                 440                 445

Thr Pro Thr Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met
    450                 455                 460

Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly
465                 470                 475                 480

Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser
                485                 490                 495

Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg
            500                 505                 510

Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala
        515                 520                 525

Leu Pro Thr Val Thr Asp Gln Gly Ala Pro Pro Val Pro Pro Thr Gly
    530                 535                 540

Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Lys Glu Ala
545                 550                 555                 560

Gln Met Pro Ala Val Ile Arg Phe
                565
```

Applicant's or agent's file reference number H. :58-1 WO	International applic. No.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 9, line 8.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution *(including postal code and country)*

23 St Machar Drive
Aberdeen AB2 1RY
Scotland, UK

Date of deposit	Accession Number
2 May 1995	NCIMB 40721

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of all designated states in which such action is possible and to the extent that it is legally permissible under the law of the designated state, it is requested that a sample of the deposited micro-organism be made available only by the issue thereof to an independent expert, in accordance with the relevant patent legislation, e.g. Rule 28(4) EPC, and generally similar provisions *mutatis mutandis* for any other designated state.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)

Applicant's or agent's file reference number H; .58-1 WO	International applicat' 'o.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 12, line 19-20.

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution *(including postal code and country)*

23 St Machar Drive
Aberdeen AB2 1RY
Scotland, UK

Date of deposit	Accession Number
2 May 1995	NCIMB 40723

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

In respect of all designated states in which such action is possible and to the extent that it is legally permissible under the law of the designated state, it is requested that a sample of the deposited micro-organism be made available only by the issue thereof to an independent expert, in accordance with the relevant patent legislation, e.g. Rule 28(4) EPC, and generally similar provisions *mutatis mutandis* for any other designated state.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Applicant's or agent's file reference number H 258-1 WO	International application No.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 17, line 18-19.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution *(including postal code and country)*

23 St Machar Drive
Aberdeen AB2 1RY
Scotland, UK

Date of deposit	Accession Number
2 May 1995	NCIMB 40722

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of all designated states in which such action is possible and to the extent that it is legally permissible under the law of the designated state, it is requested that a sample of the deposited micro-organism be made available only by the issue thereof to an independent expert, in accordance with the relevant patent legislation, e.g. Rule 28(4) EPC, and generally similar provisions *mutatis mutandis* for any other designated state.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)

CLAIMS

1. A DNA molecule comprising:
(a) a region coding for a polypeptide which is human BSSL or a biologically active variant thereof;
(b) joined to the 5′-end of said polypeptide coding region, a region coding for a signal peptide capable of directing secretion of said polypeptide from *Pichia pastoris* cells transformed with said DNA molecule; and
(c) operably-linked to said coding regions defined in (a) and (b), the methanol oxidase promoter of *Pichia pastoris* or a functionally equivalent promoter.

2. A DNA molecule according to claim 1 wherein the said signal peptide is identical to, or substantially similar to, the peptide with the amino acid sequence shown as amino acids −20 to −1 of SEQ ID NO: 2 in the Sequence Listing.

3. A DNA molecule according to claim 1 wherein the said signal peptide comprises a *Saccharomyces cerevisiae* invertase signal peptide.

4. A DNA molecule according to any one of claims 1 to 3 encoding a biologically active variant of human BSSL in which at least one of the repeat units of 11 amino acids, said repeated units being indicated in SEQ ID NO: 1, is deleted.

5. A DNA molecule according to any one of claims 1 to 4 coding for a polypeptide which has BSSL activity and an amino acid sequence which is at least 95% homologous with the sequence according to SEQ ID NO: 3 or SEQ ID NO: 4.

6. A DNA molecule according to any one of claims 1 to 5 coding for a polypeptide which has the amino acid sequence according to SEQ ID NO: 3 or SEQ ID NO: 4.

7. A vector comprising a DNA molecule according to any one of claims 1 to 6.

8. A replicable expression vector according to claim 7 which is capable of mediating expression of human BSSL, or a biologically active variant thereof, in *Pichia pastoris* cells.

9. A vector according to claim 8 which is the plasmid vector pARC 5771 (NCIMB 40721), pARC 5799 (NCIMB 40723) or pARC 5797 (NCIMB 40722).

10. Host cells of the genus *Pichia* transformed with a vector according to any one of claims 7 to 9.

11. Host cells according to claim 10 which are *Pichia pastoris* cells.

12. Host cells according to claim 11 which are *Pichia pastoris* cells of the strain GS115.

13. Host cells according to claim 12 which are PPF-1[pARC 5771] (NCIMB 40721), GS115[pARC 5799] (NCIMB 40723) or GS115[pARC 5797] (NCIMB 40722).

14. A process for the production of a polypeptide which is human BSSL, or a biologically active variant thereof, which comprises culturing host cells according to any one of claims 10 to 13 under conditions whereby said polypeptide is secreted into the culture

medium, and recovering said polypeptide from the culture medium.

ABSTRACT

The invention relates to DNA molecules, recombinant vectors and cell cultures for use in methods for expression of bile salt-stimulated lipase (BSSL) in the methylotrophic yeast *Pichia pastoris*.

COMBINED DECLARATION
AND POWER OF ATTORNEY
(Original, Design, National Stage of PCT or CIP Application)

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

DNA MOLECULES FOR EXPRESSION OF POLYPEPTIDES

the specification of which: (complete (a), (b) or (c) for type of application)

Regular or Design Application

(a) _ is attached hereto.
(b) _ was filed on as Application Serial No.
and was amended on .

PCT Filed Application Entering National Stage

(c) X was described and claimed in International Application No. PCT/SE96/00318 filed on 12 March 1996.

Acknowledgement of Review of Papers and Duty of Candor

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations §1.56.

White & Case

File No. ________

Priority Claim

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

(D) no such applications have been filed.
(E) X applications have been filed as follows:

EARLIEST FOREIGN APPLICATION(S), IF ANY FILED WITHIN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION				
Country	**Application No.**	**Date of filing**	**Date of Issue**	**Priority claimed**
India	351/MAS/95	23 March 1995		x
Sweden	9501939-4	24 May 1995		x
ALL FOREIGN APPLICATION(S), IF ANY FILED MORE THAN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION				

Continuation-in-Part

(complete this part only if this is a continuation-in-part application)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code § 112, I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.) (Filing Date) (Status - patented, pending, abandoned)

(Application Serial No.) (Filing Date) (Status - patented, pending, abandoned)

White & Case
File No. ________

Power of Attorney

As a named inventor, I hereby appoint Edward V. Filardi, Reg. No. 25,757; Nels Lippert, Reg No. 25,888; Robert B. Smith, Reg. No. 28,538; David Bender, Reg. No. 35,445; Dimitrios Drivas, Reg. No. 32,218; Cecilia O'Brien Lofters, Reg. No. 33,434, Richard J. Sterner, Reg. No. 35,372; John Scheibeler, Reg. No. 35,346; and Hans-Peter G. Hoffmann, Reg. No. 37,352 of the firm of WHITE & CASE, with offices at 1155 Avenue of the Americas, New York, New York 10036, as attorneys to prosecute this application and to transact all business in the Patent and Trademark office connected therewith.

SEND CORRESPONDENCE TO:	DIRECT TELEPHONE CALLS TO:
White & Case, Patent Department, 1155 Avenue of the Americas NEW YORK, N.Y. 10036-2787, USA	(212) 819 8200 Fax: (212) 354 8113

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR	LAST NAME DAS	FIRST NAME GOUTAM	MIDDLE NAME	
RESIDENCE AND CITIZENSHIP	CITY Bangalore	STATE OR FOREIGN COUNTRY Karnataka State	COUNTRY OF CITIZENSHIP India	
POST OFFICE ADDRESS	POST OFFICE ADDRESS Flat 11, Ahuja Apartm. 93/1, 4th Main, Malleswaram	CITY Bangalore	STATE OR COUNTRY India	ZIP CODE 560003
DATE 14 March 1996	SIGNATURE OF INVENTOR Goutam Das			
FULL NAME OF SECOND JOINT INVENTOR, IF ANY	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE AND CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE OR COUNTRY	ZIP CODE
DATE	SIGNATURE OF INVENTOR			

Check proper box(es) for any added page(s) forming a part of this declaration

/ Signature for subsequent joint inventors. Number of pages added
/ Signature by administrator(trix), executor(trix) or legal representative for deceased or incapacitated inventor. Number of pages added .
/ Signature for inventor who refuses to sign or cannot be reached by person authorized under 37 CFR 1.47. Number of pages added .

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Das, Goutam

(ii) TITLE OF INVENTION: DNA Molecules for Expression of Polypeptides

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: White & Case
(B) STREET: 1155 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: United States
(F) ZIP: 10036-2787

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/624,398
(B) FILING DATE: 04-APR-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/SE96/00318
(B) FILING DATE: 12-MAR-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: SE 9501939-4
(B) FILING DATE: 24-MAY-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Thelma A. Chen Cleland
(B) REGISTRATION NUMBER: 40,948
(C) REFERENCE/DOCKET NUMBER: 1103326-0206

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 819-8200
(B) TELEFAX: (212) 354-8113

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2428 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: mammary gland

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 82..2319
(D) OTHER INFORMATION: /product= "bile-salt-stimulated lipase"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 985..1173

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1174..1377

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1378..1575

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1576..2415

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 151..2316

(ix) FEATURE:
(A) NAME/KEY: polyA_signal
(B) LOCATION: 2397..2402

(ix) FEATURE:
(A) NAME/KEY: repeat_region
(B) LOCATION: 1756..2283

(ix) FEATURE:
(A) NAME/KEY: 5'UTR
(B) LOCATION: 1..81

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1756..1788

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1789..1821

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1822..1854

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1855..1887

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1888..1920

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1921..1953

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1954..1986

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 1987..2019

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2020..2052

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2053..2085

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2086..2118

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2119..2151

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2152..2184

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2185..2217

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2218..2250

(ix) FEATURE:
(A) NAME/KEY: repeat_unit
(B) LOCATION: 2251..2283

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Nilsson, Jeanette
Blackberg, Lars
Carlsson, Peter
Enerback, Sven
Hernell, Olle
Bjursell, Gunnar
(B) TITLE: cDNA cloning of human-milk bile-salt-stimulated lipase and evidence for its identity to pancreatic carboxylic ester hydrolase
(C) JOURNAL: Eur. J. Biochem.
(D) VOLUME: 192
(F) PAGES: 543-550
(G) DATE: Sept.-1990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ACCTTCTGTA TCAGTTAAGT GTCAAGATGG AAGGAACAGC AGTCTCAAGA TAATGCAAAG     60

AGTTTATTCA TCCAGAGGCT G ATG CTC ACC ATG GGG CGC CTG CAA CTG GTT      111
                        Met Leu Thr Met Gly Arg Leu Gln Leu Val
                        -23         -20                 -15

GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC GCG AAG CTG      159
Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala Ala Lys Leu
            -10                 -5                  1

GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA GGC GTC AAT AAG AAG      207
Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val Asn Lys Lys
    5                   10                  15

CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC AAG GGC ATC CCC TTC      255
Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly Ile Pro Phe
 20                     25                  30                  35

GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG CCA CAT CCT GGC TGG      303
Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His Pro Gly Trp
                40                  45                  50

CAA GGG ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC      351
Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala
            55                  60                  65

ACC ATC ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC      399
Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu
        70                  75                  80

AAC ATT TGG GTG CCC CAG GGC AGG AAG CAA GTC TCC CGG GAC CTG CCC      447
Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg Asp Leu Pro
    85                  90                  95

GTT ATG ATC TGG ATC TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT      495
Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His
100                 105                 110                 115

GGG GCC AAC TTC CTC AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC      543
Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala
                120                 125                 130

ACA CGC GGA AAC GTC ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC      591
Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro
            135                 140                 145

CTT GGG TTC CTC AGC ACT GGG GAC GCC AAT CTG CCA GGT AAC TAT GGC      639
Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly
        150                 155                 160

CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG AAT ATC GCG      687
Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg Asn Ile Ala
    165                 170                 175

GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG GAG TCT GCT      735
Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala
180                 185                 190                 195
```

GGA GGT GCC AGC GTC TCT CTG CAG ACC CTC TCC CCC TAC AAC AAG GGC 783
Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly
200 205 210

CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT CCC TGG 831
Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser Pro Trp
215 220 225

GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG GTG GCT GAG AAG 879
Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val Ala Glu Lys
230 235 240

GTG GGT TGC CCT GTG GGT GAT GCC GCC AGG ATG GCC CAG TGT CTG AAG 927
Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln Cys Leu Lys
245 250 255

GTT ACT GAT CCC CGA GCC CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA 975
Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala
260 265 270 275

GGC CTG GAG TAC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT GTC ATT 1023
Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro Val Ile
280 285 290

GAT GGA GAC TTC ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC AAC GCC 1071
Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala
295 300 305

GCC GAC ATC GAC TAT ATA GCA GGC ACC AAC AAC ATG GAC GGC CAC ATC 1119
Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp Gly His Ile
310 315 320

TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC AAC AAG AAA GTC 1167
Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn Lys Lys Val
325 330 335

ACG GAG GAG GAC TTC TAC AAG CTG GTC AGT GAG TTC ACA ATC ACC AAG 1215
Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr Ile Thr Lys
340 345 350 355

GGG CTC AGA GGC GCC AAG ACG ACC TTT GAT GTC TAC ACC GAG TCC TGG 1263
Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp
360 365 370

GCC CAG GAC CCA TCC CAG GAG AAT AAG AAG AAG ACT GTG GTG GAC TTT 1311
Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val Val Asp Phe
375 380 385

GAG ACC GAT GTC CTC TTC CTG GTG CCC ACC GAG ATT GCC CTA GCC CAG 1359
Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala Leu Ala Gln
390 395 400

CAC AGA GCC AAT GCC AAG AGT GCC AAG ACC TAC GCC TAC CTG TTT TCC 1407
His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser
405 410 415

CAT CCC TCT CGG ATG CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT 1455
His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His
420 425 430 435

GCA GAT GAC ATT CAG TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG 1503
Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr
440 445 450

GGC TAC CGG CCC CAA GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC 1551
Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr
455 460 465

TGG ACC AAC TTT GCC AAA ACA GGG GAC CCC AAC ATG GGC GAC TCG GCT 1599
Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly Asp Ser Ala
470 475 480

GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG GAA AAC AGC GGC TAC CTG 1647
Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser Gly Tyr Leu
485 490 495

GAG ATC ACC AAG AAG ATG GGC AGC AGC TCC ATG AAG CGG AGC CTG AGA 1695
Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg Ser Leu Arg
500 505 510 515

ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC TAT CTG GCG CTG CCC ACA 1743
Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr
520 525 530

GTG ACC GAC CAG GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC GAG 1791
Val Thr Asp Gln Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu
535 540 545

GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC GAG ACC GCC CCC GTG CCG 1839
Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro
550 555 560

CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC 1887
Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser
565 570 575

GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG 1935
Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
580 585 590 595

CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC 1983
Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp
600 605 610

TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC 2031
Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro
615 620 625

GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT 2079
Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
630 635 640

GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC 2127
Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro
645 650 655

CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG ACC CCC ACG 2175
Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Thr Pro Thr
660 665 670 675

```
GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC      2223
Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala
                680                 685                 690

CCC CCT GTG CCC CCC ACG GGT GAC TCT GAG GCT GCC CCT GTG CCC CCC      2271
Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Ala Pro Val Pro Pro
            695                 700                 705

ACA GAT GAC TCC AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT TAG      2319
Thr Asp Asp Ser Lys Glu Ala Gln Met Pro Ala Val Ile Arg Phe  *
        710                 715                 720

CGTCCCATGA GCCTTGGTAT CAAGAGGCCA CAAGAGTGGG ACCCCAGGGG CTCCCCTCCC    2379

ATCTTGAGCT CTTCCTGAAT AAAGCCTCAT ACCCCTAAAA AAAAAAAAA                2428
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 746 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Leu Thr Met Gly Arg Leu Gln Leu Val Val Leu Gly Leu Thr Cys
-23         -20                 -15                 -10

Cys Trp Ala Val Ala Ser Ala Ala Lys Leu Gly Ala Val Tyr Thr Glu
         -5                   1               5

Gly Gly Phe Val Glu Gly Val Asn Lys Lys Leu Gly Leu Leu Gly Asp
 10                  15                  20                  25

Ser Val Asp Ile Phe Lys Gly Ile Pro Phe Ala Ala Pro Thr Lys Ala
                 30                  35                  40

Leu Glu Asn Pro Gln Pro His Pro Gly Trp Gln Gly Thr Leu Lys Ala
             45                  50                  55

Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala Thr Ile Thr Gln Asp Ser
         60                  65                  70

Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu Asn Ile Trp Val Pro Gln
     75                  80                  85

Gly Arg Lys Gln Val Ser Arg Asp Leu Pro Val Met Ile Trp Ile Tyr
 90                  95                 100                 105

Gly Gly Ala Phe Leu Met Gly Ser Gly His Gly Ala Asn Phe Leu Asn
                110                 115                 120

Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn Val Ile
            125                 130                 135

Val Val Thr Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser Thr
        140                 145                 150
```

```
Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly Leu Arg Asp Gln His Met
    155                 160                 165

Ala Ile Ala Trp Val Lys Arg Asn Ile Ala Ala Phe Gly Gly Asp Pro
170                 175                 180                 185

Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala Gly Gly Ala Ser Val Ser
                190                 195                 200

Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly Leu Ile Arg Arg Ala Ile
            205                 210                 215

Ser Gln Ser Gly Val Ala Leu Ser Pro Trp Val Ile Gln Lys Asn Pro
        220                 225                 230

Leu Phe Trp Ala Lys Lys Val Ala Glu Lys Val Gly Cys Pro Val Gly
    235                 240                 245

Asp Ala Ala Arg Met Ala Gln Cys Leu Lys Val Thr Asp Pro Arg Ala
250                 255                 260                 265

Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala Gly Leu Glu Tyr Pro Met
                270                 275                 280

Leu His Tyr Val Gly Phe Val Pro Val Ile Asp Gly Asp Phe Ile Pro
            285                 290                 295

Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala Ala Asp Ile Asp Tyr Ile
        300                 305                 310

Ala Gly Thr Asn Asn Met Asp Gly His Ile Phe Ala Ser Ile Asp Met
    315                 320                 325

Pro Ala Ile Asn Lys Gly Asn Lys Lys Val Thr Glu Glu Asp Phe Tyr
330                 335                 340                 345

Lys Leu Val Ser Glu Phe Thr Ile Thr Lys Gly Leu Arg Gly Ala Lys
                350                 355                 360

Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro Ser Gln
            365                 370                 375

Glu Asn Lys Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe
        380                 385                 390

Leu Val Pro Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys
    395                 400                 405

Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser His Pro Ser Arg Met Pro
410                 415                 420                 425

Val Tyr Pro Lys Trp Val Gly Ala Asp His Ala Asp Asp Ile Gln Tyr
                430                 435                 440

Val Phe Gly Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro Gln Asp
            445                 450                 455

Arg Thr Val Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala Lys
        460                 465                 470

Thr Gly Asp Pro Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu
    475                 480                 485
```

Pro Tyr Thr Thr Glu Asn Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met
490 495 500 505

Gly Ser Ser Ser Met Lys Arg Ser Leu Arg Thr Asn Phe Leu Arg Tyr
510 515 520

Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln Glu Ala
525 530 535

Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro
540 545 550

Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly
555 560 565

Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro
570 575 580 585

Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser
590 595 600

Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
605 610 615

Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp
620 625 630

Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ala Gly Pro Pro Pro
635 640 645

Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
650 655 660 665

Asp Ser Gly Ala Pro Pro Val Thr Pro Thr Gly Asp Ser Glu Thr Ala
670 675 680

Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr
685 690 695

Gly Asp Ser Glu Ala Ala Pro Val Pro Pro Thr Asp Asp Ser Lys Glu
700 705 710

Ala Gln Met Pro Ala Val Ile Arg Phe *
715 720

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 722 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: Mammary gland

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Lys Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val
1   5   10   15

Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly
            20   25   30

Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His
        35   40   45

Pro Gly Trp Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys
    50   55   60

Leu Gln Ala Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys
65   70   75   80

Leu Tyr Leu Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg
                85   90   95

Asp Leu Pro Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly
            100   105   110

Ser Gly His Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu
        115   120   125

Glu Ile Ala Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg
    130   135   140

Val Gly Pro Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly
145   150   155   160

Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg
                165   170   175

Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly
            180   185   190

Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr
        195   200   205

Asn Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu
    210   215   220

Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val
225   230   235   240

Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln
                245   250   255

Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val
            260   265   270

Pro Leu Ala Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val
        275   280   285

Pro Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr
    290   295   300

Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp
305   310   315   320

Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn
325 330 335

Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr
340 345 350

Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr
355 360 365

Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val
370 375 380

Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala
385 390 395 400

Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr
405 410 415

Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly
420 425 430

Ala Asp His Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala
435 440 445

Thr Pro Thr Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met
450 455 460

Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly
465 470 475 480

Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser
485 490 495

Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg
500 505 510

Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala
515 520 525

Leu Pro Thr Val Thr Asp Gln Glu Ala Thr Pro Val Pro Pro Thr Gly
530 535 540

Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Thr Ala
545 550 555 560

Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr
565 570 575

Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala
580 585 590

Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro
595 600 605

Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly
610 615 620

Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro
625 630 635 640

Pro Thr Gly Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp Ser
645 650 655

Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
660 665 670

Thr Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp
675 680 685

Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Ala Pro
690 695 700

Val Pro Pro Thr Asp Asp Ser Lys Glu Ala Gln Met Pro Ala Val Ile
705 710 715 720

Arg Phe

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 568 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: Mammary gland

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..568
(D) OTHER INFORMATION: /label= Variant_C

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Hansson, Lennart
Blackberg, Lars
Edlund, Michael
Lundberg, Lennart
Stromqvist, Mats
Hernell, Olle
(B) TITLE: Recombinant Human Milk Bile Salt-stimulated Lipase
(C) JOURNAL: J. Biol. Chem.
(D) VOLUME: 268
(E) ISSUE: 35
(F) PAGES: 26692-26698
(G) DATE: Dec. 15-1993

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ala Lys Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val
1 5 10 15

Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly
20 25 30

Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His
35 40 45

```
Pro Gly Trp Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys
    50                  55                  60

Leu Gln Ala Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys
65                  70                  75                   80

Leu Tyr Leu Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg
                85                  90                  95

Asp Leu Pro Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly
            100                 105                 110

Ser Gly His Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu
        115                 120                 125

Glu Ile Ala Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg
    130                 135                 140

Val Gly Pro Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly
145                 150                 155                 160

Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg
                165                 170                 175

Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly
            180                 185                 190

Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr
        195                 200                 205

Asn Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu
    210                 215                 220

Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val
225                 230                 235                 240

Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln
                245                 250                 255

Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val
            260                 265                 270

Pro Leu Ala Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val
        275                 280                 285

Pro Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr
    290                 295                 300

Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp
305                 310                 315                 320

Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn
                325                 330                 335

Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr
            340                 345                 350

Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr
        355                 360                 365

Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val
    370                 375                 380
```

Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala
385 390 395 400

Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr
405 410 415

Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly
420 425 430

Ala Asp His Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala
435 440 445

Thr Pro Thr Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met
450 455 460

Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly
465 470 475 480

Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser
485 490 495

Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg
500 505 510

Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala
515 520 525

Leu Pro Thr Val Thr Asp Gln Gly Ala Pro Pro Val Pro Pro Thr Gly
530 535 540

Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Lys Glu Ala
545 550 555 560

Gln Met Pro Ala Val Ile Arg Phe
565